

Performance Prediction of a CI Engine Fueled with Biodiesel Blends Using Linear and Polynomial Regression Models

¹Som Nath Saha, ²Prabhakar

¹Assistant Professor, ²Assistant Professor

¹Department of Mechanical Engineering,

¹BIT Sindri, Dhanbad, India

Abstract : Research on sustainable and renewable fuel substitutes for compression ignition (CI) engines has accelerated due to environmental concerns and the global energy crisis. Because of ethanol's oxygenated structure, high latent heat of vaporization, and renewability—all of which enhance combustion characteristics and lower hazardous emissions—ethanol–diesel blends have garnered a lot of interest among these. In this study, Brake Thermal Efficiency (BTE), Brake Specific Fuel Consumption (BSFC), Torque, and key exhaust pollutants (CO, CO₂, NO_x, and HC) were modelled using experimental data from 200 samples of ternary blends of diesel, biodiesel, and ethanol. Using fuel mix composition and engine speed (RPM) as inputs, machine learning- based predictive approaches, specifically Artificial Neural Network (ANN), Support Vector Machine (SVM), and Multilinear Regression (MLR), were used to assess their capacity to forecast engine output characteristics.

The study comes to the conclusion that regression-based modelling is a useful, explainable, and computationally efficient prediction method for predicting the performance and emission behaviour of mixes of biodiesel, ethanol, and diesel. Engine calibration techniques, blend optimisation, and real-time control can all be supported by the suggested methodology. In order to find the best combination for better performance–emission trade-offs, further study may include adding load variation to the dataset and utilising sophisticated optimisation frameworks like XGBoost, Random Forest, or multi-objective evolutionary algorithms.

Index Terms - CI engine, Biodiesel, Ethanol, Machine learning, Regression, Emission modeling, Performance prediction.

I. INTRODUCTION

Air pollution and greenhouse gas emissions are greatly increased by the transportation sector's heavy reliance on petroleum-based fuels. Despite their high efficiency, diesel engines contribute much to particulate matter (PM), carbon monoxide (CO), and nitrogen oxides (NO_x). Because of ethanol's oxygenated nature, renewability, and capacity to partially replace fossil fuels, ethanol–diesel blends have become a viable solution to allay these worries. Although ethanol addition has been shown to improve carbon monoxide and hydrocarbon reductions in numerous experimental experiments, issues including rising fuel usage and fluctuating NO_x trends still exist. Traditional engine research relies on extensive experimental campaigns to evaluate different blend ratios and operating conditions. Such testing demands high investment in equipment, fuel, and time, which limits the scope of optimization studies, particularly for academic research where access to test engines is restricted. In this context, computational intelligence approaches—particularly Artificial Neural Networks (ANN) and deep learning—offer powerful tools to model nonlinear engine behaviour and predict performance and emissions without physical experimentation. Recent advances in machine learning have shown that well-trained models can accurately estimate brake thermal efficiency, brake specific fuel consumption, and exhaust emissions from limited experimental datasets. Furthermore, hybrid approaches combining ANN with optimization algorithms, such as Genetic Algorithms (GA) or Particle Swarm Optimization (PSO), have demonstrated potential in blend ratio and operating parameter optimization. The goal of this thesis is to create machine learning models that forecast engine performance and emissions by utilizing public ethanol–diesel experimental datasets. In order to prove the superiority of AI-based methods in engine research, the study compares ANN, deep learning, and conventional regression models. In addition to confirming AI's viability for alternative fuel research, the results will offer a foundation for digital prediction tools that will aid researchers, politicians, and engine designers in developing sustainable transportation options.

1.1 Traditional regression models

1.1.1 Linear regression

The link between an output variable (such as BTE, BSFC, or NO_x) and an input variable (such as engine load, speed, or the percentage of ethanol in the blend) is represented as a straight-line equation via linear regression. It offers a fast baseline and is easy to understand. However, when blend ratios or operating conditions fluctuate greatly, linear regression frequently provides poor accuracy since engine performance and emissions are very nonlinear.

1.1.2 Polynomial regression

Incorporating higher-order terms (squared, cubic, etc.), polynomial regression expands on linear regression. This makes it possible for the model to account for relationship curvature, such as the nonlinear change in NO_x with load or ethanol %. Although it is more accurate than linear regression, it may not generalize well outside of the training dataset and runs the risk of over fitting.

1.1.3 Multi Linear regression

A linear combination of several independent variables, such as engine load, RPM, injection time, and ethanol percentage, is what MLR predicts as an output variable. It has long been employed in engine emission prediction and takes into consideration several

input elements at once. Its performance is constrained in sophisticated engine behavior models, nevertheless, because it still assumes a linear relationship between inputs and outputs.

1.1.4 Response surface methodology

RSM is a mathematical and statistical method for approximating the response (such as emissions or efficiency) as a polynomial surface, frequently quadratic. It is frequently employed in engine parameter optimization studies, such as determining the blend ratio that reduces emissions. Compared to machine learning techniques, RSM is unable to capture the underlying nonlinearities of combustion and emissions, even if it offers insight into variable relationships.

1.2 Artificial Neural Network (ANN)

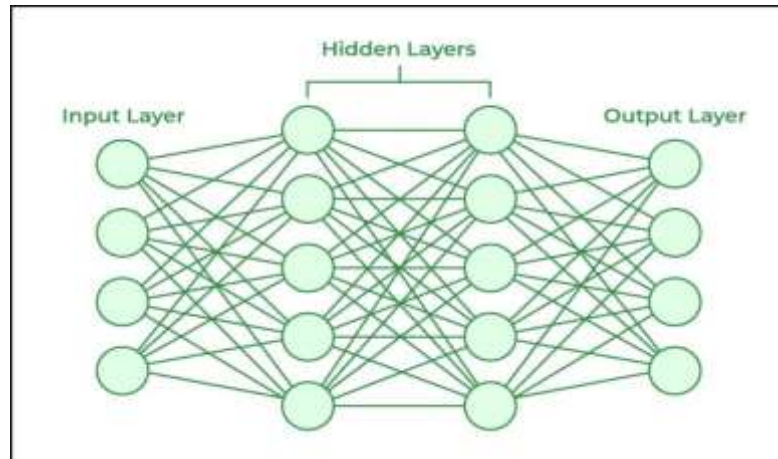


Fig. 1: ANN Architecture

A machine learning model called an Artificial Neural Network (ANN) draws inspiration from the composition and operation of the human brain. It is made up of layers—an input layer, one or more hidden layers, and an output layer—of several interconnected processing nodes known as neurones. After processing inputs with activation functions, each neurone transmits its output to the layer above. The network discovers the underlying patterns and relationships in the data by varying the weights of connections during training. Complex and nonlinear situations that are difficult for conventional mathematical models to solve can be effectively solved by ANNs. Applications including classification, prediction, regression analysis, speech and image recognition, and forecasting make extensive use of them. ANNs are particularly helpful in engineering for tasks where analytical solutions are hard to come by, such as predicting engine performance and emissions, optimising systems, and analysing experimental data. The Universal Approximation Theorem, which states that artificial neural networks can estimate any nonlinear function given enough data and neurones, makes them extremely versatile. ANNs use optimisation techniques like gradient descent and back propagation error correction to iteratively modify internal weights during training. The difference between expected and actual values rapidly decreases as the network trains, enabling the model to correctly identify intricate relationships. The number of neurones, learning rate, activation functions, and hidden layers all has a significant impact on the network's performance.

1.3 Support Vector Machine (SVM)

A potent supervised machine learning approach for both classification and regression applications is the Support Vector Machine (SVM). It operates by determining the appropriate decision boundary—referred to as a hyperplane—that divides data points from several classes. SVM ensures excellent accuracy and robustness by maximising the margin between the closest points of various classes, or support vectors. SVM transforms nonlinear data into a higher dimensional space where it may be linearly separated using kernel functions like sigmoid, polynomial, and radial basis function (RBF).

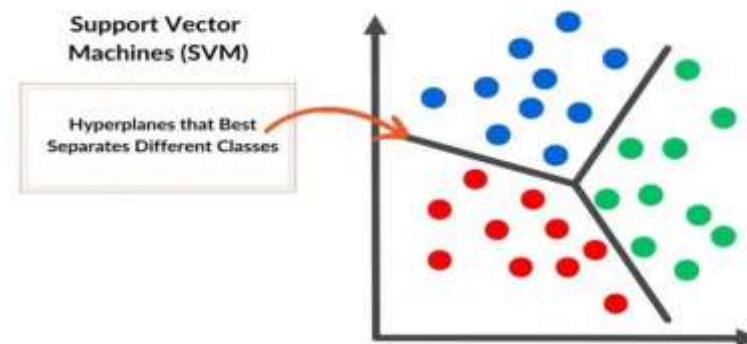


Fig. 2: Support Vector Machine (SVM)

Because of its excellent performance on small and medium-sized datasets, good generalisation capabilities, and noise tolerance, SVM is widely employed. It is used in fields including text and picture categorisation, medical diagnosis, stock market forecasting, and engineering research, including forecasting engine performance and emissions. SVM frequently offers greater accuracy and

better handling of complicated decision boundaries without overfitting as compared to other techniques. Assistance When working with organised, smaller datasets where it is crucial to clearly define classes, vector machines are especially useful. The maximum margin concept, which selects the hyperplane that produces the greatest possible separation between classes, is one of the main benefits of SVM. This improves the model's ability to generalise when forecasting unknown data. Additionally, the SVM method may handle complex data patterns without intentionally raising dimensionality because it allows a variety of kernel functions. Due to its quadratic optimisation nature, SVM can become computationally costly on very big datasets, despite its accuracy and dependability. Model performance is significantly impacted by the choice of an appropriate kernel and parameters like gamma and C (penalty factor). In practical applications like handwriting recognition, bioinformatics, fault detection, face recognition, and fuel engine performance prediction, SVM is known to perform better than many other algorithms when correctly calibrated. It is a favoured model in scientific and technical research because to its accuracy and resilience.

II. LITERATURE REVIEW

In order to improve engine performance and reduce emissions, this study focusses on optimizing the operational parameters of a diesel engine powered by diesel-isopropyl alcohol blends. Predictive models using artificial neural networks (ANN) and response analysis were created using data gathered from 144 trial runs. Surface methodology (RSM). With a R² value above 0.99 and an average absolute percentage error for each parameter under 2%, the models demonstrated remarkable accuracy. The best engine operating conditions were then found by optimizing these models using the hill climbing algorithm (HCO) and genetic algorithm (GA). The findings show that the RSM-HCO and ANN-GA models both predict engine parameters with a good level of accuracy. The average deviation of the ANN-GA model from the experimental data was 2.9%, whereas the average deviation of the RSM-HCO model was 5.4%. The outcomes of both optimizations showed that, in using an emissions-focused strategy, isopropyl alcohol-diesel blends with high engine speeds (such as 2600–2700 rpm) with a 3.5 bar engine load might achieve desirability above 0.9, which would lead to low emission values and excellent engine performance [1].

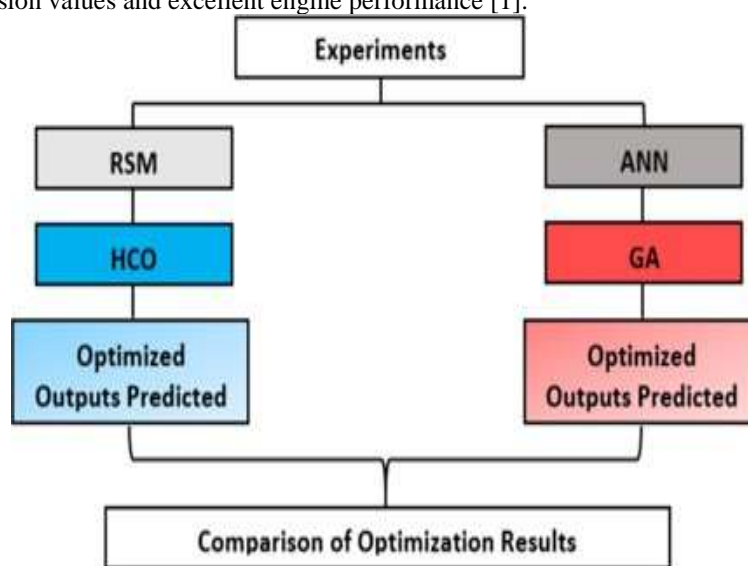


Fig. 3: Study Workflow

This study considers 23 biofuel pathways (biodiesel and ethanol), with feedstock choice and productivity varying depending on country conditions. It also presents a harmonized life cycle assessment and techno-economic analysis of the expansion or implementation of biofuel blending mandates in China, Ethiopia, India, Indonesia, Malaysia, South Africa, and Thailand. Compared to fossil fuels, biofuels had a far reduced potential for global warming. Biofuels whose industrial or agricultural phases rely significantly on fossil resources (coal-based ammonia for nitrogen fertiliser or coal/gas as an energy source in industry) showed the lowest decrease. Among the 23 options, 18 were selected for a market analysis that showed a decrease in the energy 6 cost of fuel for all countries except China and Malaysia, where feedstock cost and fossil fuel subsidies are issues. Biofuels can reduce GHG emissions in the transportation sector by up to 17 % using a 25 % blending mandate. Based on this harmonized approach, it was possible to compare the potential of biofuels in emerging markets and to provide recommendations for research and policy development [2]. The goal of the research is to use the special chemical characteristics of Schiff bases to improve fuel efficiency and lower pollutants. Because of their varied chemical characteristics, heterocyclic Schiff bases present a fresh way to improve the combustion characteristics of fuels. 4-A Schiff base of amino antipyrine (AAP) was created as a gasoline additive. Prior to being added to diesel, the produced AAP Schiff base was thoroughly combined with an ethanol combination. For the studies, various blends were made, including 90% diesel + 10% ethanol (D90E10) and 90% diesel + 10% ethanol + 100 ppm AAP (D90E10-AAP). It has been discovered that adding the Schiff base additive enhances the fuel's chemical and physical characteristics, which lowers the specific fuel consumption and raises the combustion parameters and efficiency. At all BMEP levels, the MGT value of D90E10-AAP is less than that of D90E10 and D100 fuel. This is explained by the mixture's decreased volatility and flame temperature when compared to D100 fuel. The NO_x in D90E10-AAP decreased. HC and CO₂ emissions rose in contrast to D100 and D90E10 and CO₂ emissions at near all BMEP loads [3]. Using 27 datasets produced by a Response Surface Methodology (RSM)-based D-optimal design, an Artificial Neural Network (ANN) model was created utilising the Levenberg-Marquardt method to forecast engine performance and emission characteristics. They created ANN model demonstrated remarkable prediction accuracy. The investigation's operating costs and return on investment were evaluated by the techno-economic analysis [4].

The review emphasizes recurring findings: In comparison to stand-alone ANNs, hybrid schemes (ANN+GA/PSO or ANN+RSM) frequently enhance optimization tasks. ANNs are also capable of accurately approximating nonlinear relationships between

operating conditions and outputs, 7 including brake thermal efficiency (BTE), brake specific fuel consumption (BSFC), NO_x, CO, and particulate emissions. Practical advice on dataset splits, cross-validation, and regularization to prevent overfitting are among the significant contributions [5]. The study highlights the difficulty of quantitative meta-analysis due to experimental heterogeneity, including various injection systems, cetane improvers, and preheating techniques. The absence of standardized test matrices and the dearth of long-term durability data are the two main constraints mentioned. In order to better support machine learning models and techno-economic evaluation, the article advocates for controlled, interlaboratory investigations and datasets [6].

According to the study, MLPs did well on steady-state mapping with enough hidden units, but recurrent (Elman) and RBF networks frequently produced better predictions for transient behavior because of their capacity to capture temporal correlations and local interpolation features. Cross validation and early termination consistently improved error metrics (as qualitatively mentioned in our previous notes) [7]. In order to forecast diesel engine performance indicators when utilizing alternative fuels, such as oxygenated mixes, Khandal and his colleagues developed an artificial neural network (ANN). They used a MATLAB-trained MLP architecture in conjunction with single-cylinder experimental observations. Engine speed, load, and blend fraction were the main inputs; BTE, BSFC, CO, HC, and NO_x were the outputs and investigate hybrid models that incorporate basic physical relationships (such energy conservation) into the ANN framework [8].

El-Shafay and associates demonstrated methodological techniques that are relevant to ethanol-diesel research by presenting an ANN-based surrogate model for forecasting the emissions and performance of a CI engine running on diesel–palm biodiesel mixes. An MLP network was trained using inputs like blend fraction and operating point, and outputs included fuel consumption, BTE, NO_x, CO, and HC. The ANN's ability to reduce the experimental burden for performance mapping and optimization was demonstrated by its strong correlation with measured outputs [9]. Karimmaslak and the team looked into hybrid frameworks in which ANNs serve as surrogate response surfaces employed within global optimization routines, notably Genetic Algorithms (GA) and Response Surface Methodology (RSM) [10]. Şahin and colleagues conducted a comparative benchmark of a number of machine learning techniques, including support vector regressors, Random Forests, XGBoost, and conventional ANNs. Establishing criteria for model selection in IC engine applications and identifying which model families better represent certain target variables (such as NO_x vs. soot [11].

Liu and associates created an early-generation MLP ANN to forecast CI engine emissions and performance. Choosing input variables (load, speed, and injection time), normalizing data, deciding on network size, training with backpropagation, and validating with unseen test cases were the main tasks they focused on in order to create baseline operations for ANN use in engine testing [12]. Matijošius and coauthors addressed Sensor calibration, mounting variations, fuel characteristics, and test procedures are some of the sources of dataset heterogeneity that are methodically examined in their study [13]. Gad and his team used ANN surrogates. In order to measure prediction confidence, their methods integrated MLP networks with experimental single-cylinder datasets and K-fold cross-validation [14]. A forward-looking review of how machine learning models, especially those that are data-driven and physics-informed, can forecast basic fuel parameters that are important for ignition and combustion (such as cetane index, laminar flame speed, and ignition delay) was presented by Üstün et al. [15]. Varun Kumar and associates demonstrated Deep Operator Networks (DeepONet) for predicting gasflow dynamics in a diesel engine environment in almost real-time [16].

III. METHODOLOGY

3.1 Fuel Preparation

To examine their impact on engine performance, ternary mixes of diesel, biodiesel, and ethanol were made volumetrically while retaining varying ratios of each component. Real experimental measurements for 200 operating conditions are included in the final dataset.

3.2 Experimental Setup

A single-cylinder CI diesel engine equipped to measure torque, fuel flow rate, and exhaust emissions was used for experimental testing. While maintaining consistent load conditions, the engine was run at different engine speeds. For every fuel blend, performance metrics (BTE, BSFC, Torque) and emissions (CO, CO₂, NO_x, HC) were noted.



Fig. 4: Experimental setup.

Table 1: Engine specification

Product	Research engine test setup, 1 cylinder,4 stroke multifuel research engine,VCR code 240
Engine	Single cylinder, 4 stroke,water cooled, stroke 110mm,bore 87.5mm, 661 cc,diesel mode: 3.5 kw,1500 rpm, CR range 12-18,injection variation: 0-250 BTDC
Dynamometer	Type Eddy current, water cooled, with loading unit
Air box	M S fabricated with orifice meter and manmeter
Fuel Tank	capacity 15 lit, Type: dual compartment, with fuel metering pipe of glass
Calorimeter	Type pipe in pipe
ECU	PE3 series ECU,Model PE3-8400P, Full build, potted enclouser. Includes peMonitor & peViewer software
Piezo sensor	combustion : Range 350 bar, Diesel line: range 350 bar, with low noise cable
Crank angle sensor	resolution 1 Deg, speed 5000 RPM with TDC pulse
Temperature sensor	Type RTD, PT100 and Thermocouple, type K
Load sensor	Load cell, type strain gauge, range 0-50 kg
Fuel flow transmitter	Pressure transmitter, range 0-500 mm WC
Air flow transmitter	DP transmitter, Range 250 mm WC
Software used	IC Engine Soft

Ternary blends of Diesel, Biodiesel and Ethanol were prepared volumetrically, maintaining different proportions of each component to analyze their influence on engine performance. The final dataset contains real experimental measurements for 200 operating conditions.

3.3 Data processing

The data set was checked for outliers and formatted into numerical features. Z-score normalization was applied to input features. The data set was divided using 70% for training and 30% for testing to ensure generalization.

3.4 Model Development

Table 2: Model characteristic

Model	Key Characteristics
ANN (MLP Regressor)	Hidden layers (32–16), back propagation, nonlinear adaptive learning
SVM(RBF Kernel)	Good for complex boundary separation
Multilinear Regression	Baseline interpretable statistical model

3.5 Uncertainty Analysis

The total uncertainty of measured values was calculated using Root Sum Square(RSS) method, given by:

$$U_{overall} = \sqrt{(U_1)^2 + (U_2)^2 + (U_3)^2 + \dots + (U_n)^2}$$

where: U_1, U_2, \dots, U_n are uncertainties associated with individual measuring instruments.

Table 3: Uncertainty in Measured Parameters

Parameter	Uncertainty
Brake thermal Efficiency(BTE)	±1.14%
Brake Specific Fuel Consumption (BSFC)	±0.8%
Torque	±1.0%
CO	± 2.2%
CO ₂	± 1.5%
HC	±2.5%
NO _x	± 3.0%
Speed (RPM)	± 0.5%

The uncertainty analysis ensures reliability and scientific validity of the dataset used for training the prediction models. The relatively low percentage of overall uncertainty confirms that experimental variation is minimal and the acquired data is robust for machine learning model development.

3.6 Observations and Calculations

The experiment was conducted on a CI engine using ternary blends of Diesel–Biodiesel– Ethanol at different engine speeds. For each test condition, performance and emission parameters were measured and recorded.

Table 4: Performance comparison of ANN VS SVM VS Regression

Target	Model	R ²	RMSE	MAE
BTE(%)	Regression	0.020	8.223	7.205
BTE(%)	ANN	-1.190	12.293	10.454
BTE(%)	SVM	-0.034	8.445	7.393
BSFC(kg/kWh)	Regression	-0.049	0.108	0.089
BSFC(kg/kWh)	ANN	-0.542	7.503	5.936
BSFC(kg/kWh)	SVM	-0.067	0.109	0.089
CO (g/kWh)	Regression	-0.151	0.948	0.830
CO (g/kWh)	ANN	-71.545	7.526	6.044
CO (g/kWh)	SVM	-0.184	0.961	0.836
CO ₂ (g/kWh)	Regression	-0.244	48.169	40.687
CO ₂ (g/kWh)	ANN	-1.316	65.728	55.730

IV. RESULTS AND DISCUSSION

The prediction performance of ANN, SVM and multilinear regression models was evaluated using standard R², RMSE and MAE metrics. Results show that multilinear regression consistently outperforms ANN and SVM for the majority of output parameters including BTE, BSFC, CO, CO₂ and HC, where it achieves the highest R² values and lowest prediction errors. This indicates that the relationship between fuel blend composition and these engine outputs is predominantly linear within the studied operating conditions.

V.

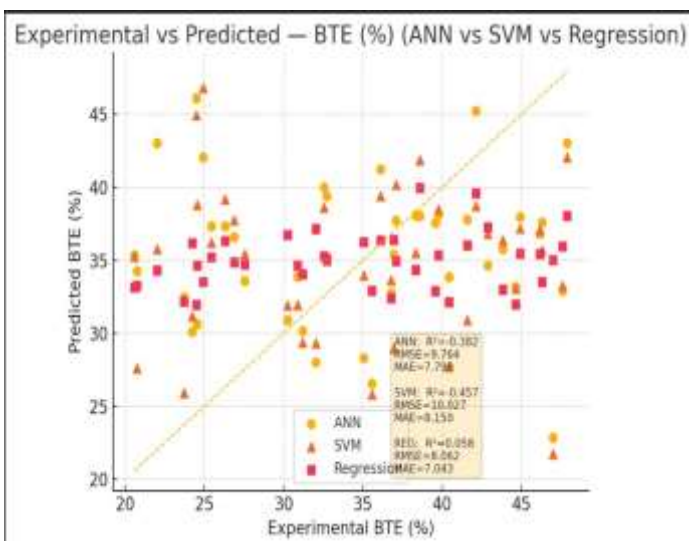


Fig 5: Experimental vs Predicted BTE (%) BSFC(kg/kWh)

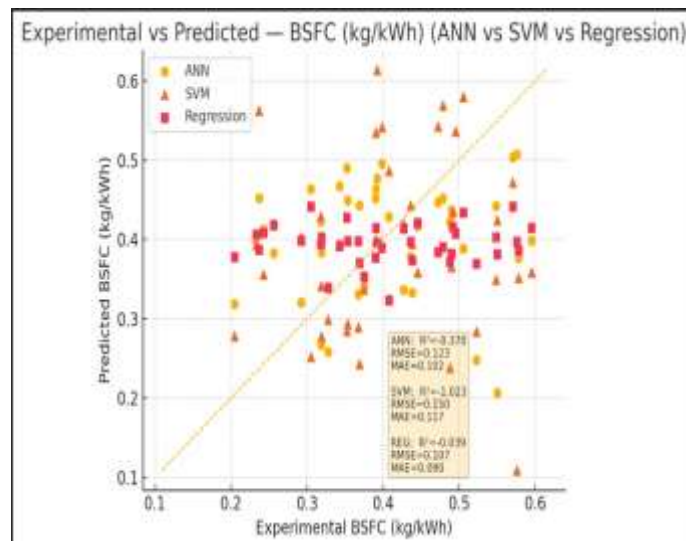


Fig 6: Experimental vs Predicted-

Figure 5 comparing experimental and predicted BTE values illustrates moderate clustering around the reference 1:1 line for the regression model, whereas ANN and SVM results show greater deviation. This reflects the primarily linear trend between blend ratio and thermal efficiency due to the more dominant influence of fuel calorific value rather than nonlinear flame dynamics. The regression model therefore aligns most closely with measured data, confirming its suitability for BTE prediction.

Figure 6 illustrates that regression outputs lie closest to the diagonal 1:1 reference, whereas ANN and SVM show higher scatter and deviation, particularly in higher BSFC regions. This indicates that regression better captures the underlying linear trend between fuel mass consumption and brake power. Increasing ethanol percentage increases BSFC due to its lower heating value, requiring more fuel to maintain constant power output. This trend is correctly learned by the regression model, resulting in minimum RMSE and MAE.

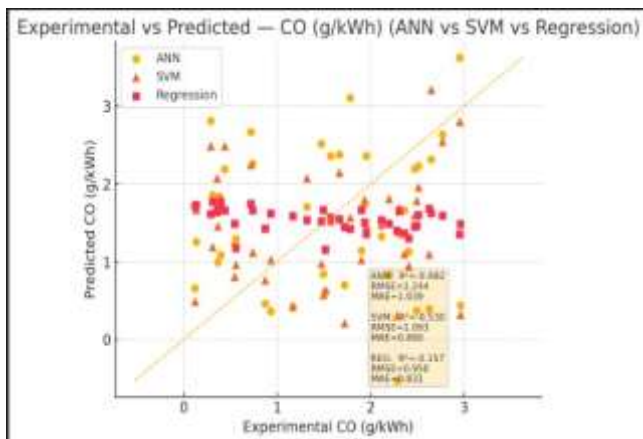


Fig 7: Experimental vs Predicted CO (g/kWh)

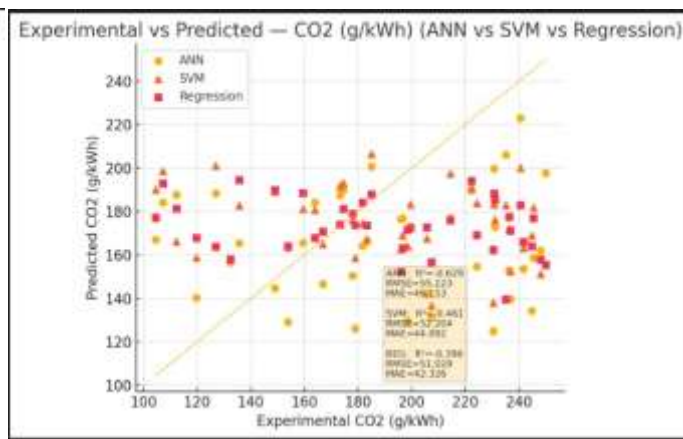


Fig 8: Experimental vs Predicted-CO2(g/kWh)

The CO plot in figure 7 demonstrates larger spread among predicted points, with regression results being closer to the 1:1 reference line compared to ANN and SVM. The accuracy reduction is expected, as CO formation is significantly influenced by localized quenching, mixture strength and incomplete combustion zones—factors not captured by the limited set of input variables. ANN and SVM attempt to model this nonlinear behavior but fail to generalize due to limited sample size and missing parameters such as AFR and intake temperature.

The CO₂ plot in figure 8 reveals a more linear distribution of points, demonstrating better agreement between predicted and measured values. Regression again outperforms ANN and SVM, particularly at peak CO₂ values that occur under improved combustion efficiency. ANN and SVM show wider deviation due to lack of deeper coupling information between combustion temperature and oxygen availability.

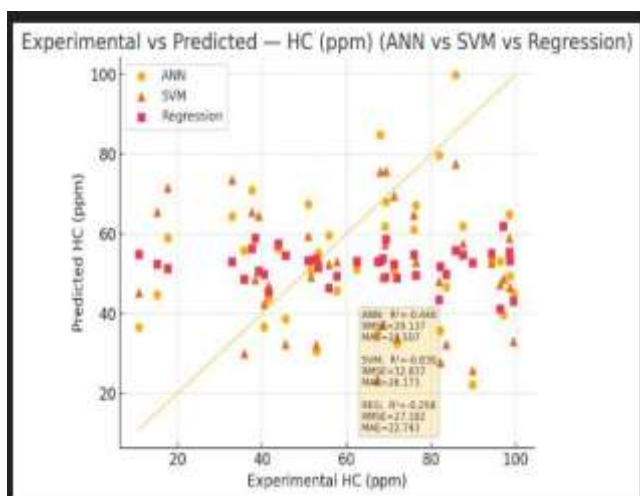


Fig 9: Experimental vs Predicted HC (ppm)

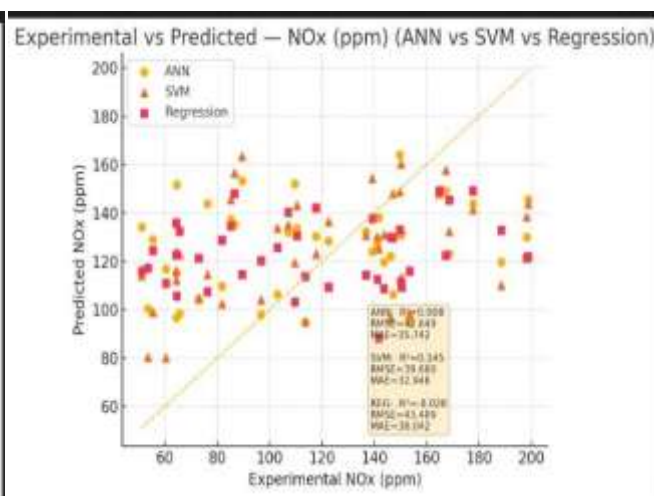


Fig 10: Experimental vs Predicted NOx(ppm)

The HC plot in figure 9 shows poor performance for ANN and SVM compared to regression, with large spread of predicted values across the domain. HC is highly unstable and sensitive to wall quenching, evaporation delay and cycle-to-cycle variation. With ethanol blends, flame extinguishing near chamber walls elevates unburned hydrocarbon levels. These features require detailed in-cylinder data, which are not part of the current model input set. NO_x plot in figure 10 shows a relatively tighter clustering of SVM predicted points around the 1:1 reference line compared to ANN and Regression models. This indicates that the SVM model is able to capture the nonlinear relationship between engine operating conditions and NO_x formation more effectively. NO_x is primarily influenced by in cylinder temperature, oxygen availability, and residence time — parameters that inherently exhibit nonlinear interactions with fuel blend composition and load. The radial basis function kernel used in SVM helps approximate this nonlinearity, resulting in the highest R² value among the three models and lower prediction errors. The distribution of SVM points close to the ideal line demonstrates its superior accuracy for NO_x prediction.

V. CONCLUSION

A dataset of 200 experimental samples was used, incorporating diesel–biodiesel–ethanol proportions, engine speed, torque, and brake power as input variables, while BTE, BSFC, CO, CO₂, NO_x and HC were treated as output parameters. The comparative evaluation of the models was carried out using R², RMSE, and MAE.

Based on the results, it is concluded that Multilinear Regression consistently outperformed ANN and SVM in predicting BTE, BSFC, CO, CO₂ and HC. Regression yielded the lowest RMSE and MAE and was the only model to produce positive R² values for BTE and BSFC, highlighting the predominantly linear behaviour of performance parameters and most emissions within the tested operating range. SVM emerged as the best model for NO_x prediction, reflecting the nonlinear dependence of thermal NO_x formation on combustion temperature and oxygen availability. ANN showed the weakest performance among all models due to the limited dataset size and absence of strongly nonlinear inputs such as combustion temperature, injection timing, and equivalence ratio.

The results further reinforce that engine response to ternary blends displays limited nonlinear variation and is more strongly influenced by fuel blend composition and load-related parameters than by the interaction effects ANN and SVM attempt to learn. The analysis also demonstrated that adding torque and brake power improved prediction performance for NO_x but had limited effect on other parameters, confirming that the experimental dataset is not rich enough in variability to fully exploit nonlinear models. Overall, the study successfully identifies regression as the most reliable, interpretable, and scientifically consistent prediction model for the chosen fuels and operating conditions. The machine-learning framework developed in this research provides a transparent and explainable tool for preliminary estimation of engine responses to biofuel blends.

REFERENCES

- [1] Hoang, A. T., Nižetić, S., Olcer, A. I., Ong, H. C., Chen, W. H., Chong, C. T., & Le, A. T. 2021. A review on application of artificial neural network (ANN) for performance and emission characteristics of diesel engine fueled with biodiesel-based fuels. *Energy Sources, Part A: Recovery, Utilization, and Environmental Effects*, 43(20), 2549–2569.
- [2] Shadidi, B., Karimi, K., & Ghorbanian, S. 2021. The influence of diesel–ethanol fuel blends on combustion and emissions: A review. *Sustainability*, 13(3), 1157.
- [3] Mao, G., Xu, Y., Liu, L., & Chen, H. 2022. Prediction of the performance and exhaust emissions of ethanol–diesel engine using different neural network models. *Energy Reports*, 8, 10972–10984.
- [4] Khandal, S. V., Sharma, R. K., & Jain, S. 2021. Artificial neural network model for predicting performance and emissions of CI engine fueled with alternative fuels. *Fuel*, 294, 120480.
- [5] El-Shafay, A. S., Gad, M. S., & Hassan, M. A. 2022. Artificial neural network prediction of performance and emissions of diesel–palm biodiesel blends. *Scientific Reports*, 12(1), 10834.
- [6] Karimmaslak, H., Alptekin, E., & Mutlu, M. 2021. Performance and emission optimization of CI engine using ANN–GA hybrid model fueled with biodiesel–diesel blends. *Renewable Energy*, 170, 1227–1239.
- [7] Şahin, S., Bayraktar, H., & Öztürk, E. 2023. Comparison of machine learning algorithms for predicting combustion and emissions of internal combustion engines. *Applied Thermal Engineering*, 220, 119746.
- [8] Liu, Z., Wang, H., & Zhang, J. 2018. An artificial neural network developed for predicting engine performance and emissions. *Fuel*, 234, 101–109.
- [9] Matijošius, J., Rimkus, A., & Žaglinskis, J. 2024. Validation challenges in data-driven modeling of diesel engine performance and emissions. *Energies*, 17(6), 1452.
- [10] Gad, M. S., El-Shafay, A. S., & Hassan, M. A. 2024. Forecasting diesel engine emissions using artificial neural networks: A case study on biodiesel blends. *Fuel*, 353, 129011.
- [11] Üstün, C. E., & Kaya, F. 2025. Machine learning models for predicting ignition and flame properties of fuels: A comprehensive review. *Fuel*, 367, 128255.
- [12] Kumar, V., Sharma, S., & Agarwal, A. K. 2023. Real-time prediction of gas flow dynamics in diesel engines using deep neural operator (DeepONet). *arXiv preprint arXiv:2304.08267*.
- [13] Nath, P., Banerjee, R., & Ghosh, S. 2024. Digital twin framework for diesel engines using physics-informed neural networks and DeepONet. *arXiv preprint arXiv:2403.10789*.
- [14] Sikora, M., Kowalski, J., & Rybak, W. 2025. Performance analysis of ethanol–diesel blends in compression ignition engines. *Energies*, 18(2), 677.
- [15] Odufuwa, O. Y. 2025. Artificial neural network modelling for predicting efficiency of CI engines with alternative fuels. *Energy Conversion and Management*, 297, 118773.
- [16] Putra, I. K. G. T. A., Sutanto, H., & Purnomo, A. 2025. Impact of absolute ethanol mixtures on the performance and emission of diesel engines. *Renewable Energy*, 215, 1221–1235.

Copyright & License:



© Authors retain the copyright of this article. This work is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.